Reprinted from "Flatness, Roughness, and Discrete Defect Characterization for Computer Disks, Wafers, and Flat Panel Displays," John C. Stover, Editor, Proc. SPIE 2862, 12-17 (1996)

Proposed methodology for characterization of microroughness-induced optical scatter instrumentation

Thomas A. Germer and Clara C. Asmail

Optical Technology Division
National Institute of Standards and Technology
Gaithersburg, Maryland 20899

ABSTRACT

Measurements of optical scatter are often employed in production line diagnostics for surface roughness of silicon wafers. However, the geometry of the optical scatter instrumentation lacks universal standardization, making it difficult to compare values obtained by instruments made by different manufacturers. The bidirectional reflectance distribution function (BRDF), on the other hand, is a well-defined quantity, and under conditions usually met with bare silicon wafers, can be related to the power spectral density (PSD) of the surface roughness. In this paper, we present an approach for characterizing low level optical scatter instrumentation using a spatial frequency response function. Methods for calculating or measuring the response function are presented. Limitations to the validity of the spatial frequency response function are also discussed.

Keywords: surface roughness, BRDF, silicon wafers, haze, optical scatter

1. INTRODUCTION

As the critical dimensions of integrated circuit device structures become smaller the tolerance for surface microroughness becomes tighter. Scanning surface inspection systems (SSIS) are currently employed to measure levels of optical scatter and to detect particulate contamination. These instruments rely upon integrated light scatter, and treat the background signal on the wafer as resulting from surface microroughness and localized signals as resulting from particulate contamination.

Recent efforts have focused on determining a means for characterizing optical scatter instrumentation. An important step was made when it was recognized that different instruments measure roughness on different spatial frequency scales, which results from measurements being carried out with differing optical geometries, wavelengths, polarizations, etc.[1] In this paper, we will expand upon the notion of a bandwidth, by defining the spatial frequency response function, and provide a means for calculating or measuring that function.[2] Knowledge of this function will allow absolute calibration of low-level optical scatter instruments by enabling an appropriate integration of bidirectional reflectance distribution function (BRDF) data that could then be correlated to SSIS data. Furthermore, manufacturers of optical scatter intrumentation will have a well-defined methodology for specifying their products.

2. THEORY

In order to make a comparison between BRDF and any optical scatter measurement, it is neccesary to ascertain the geometry of the scattered light collection system for the instrument. Most instruments report a bandwidth by specifying minimum and maximum spatial frequencies for which their instrument is sensitive, thus making an implicit assumption that the response is uniform over that bandwidth. It is the purpose of this section to determine the methodology for converting measured BRDF to integrated scatter and to introduce the concept of the spatial frequency response function. This spatial frequency response function will serve as an improvement upon the notion of the instrument bandwidth.

The sample geometry is illustrated in Fig. 1. We begin by assuming that a particular scattered light collection system has light incident on a sample at an angle θ_i in the xz plane in a polarization state represented by q_i. Light is scattered by the sample into different angles (θ_s, ϕ_s) with polarization q_s. The collection system is assumed to

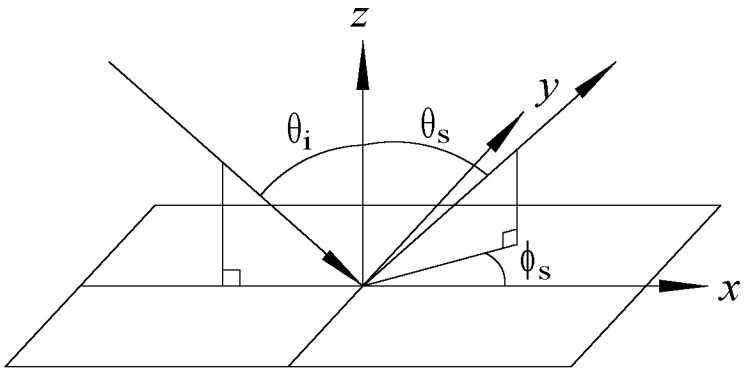

Figure 1 The geometry used in this paper.

have an angle- and polarization-dependent detection efficiency $\epsilon(\theta_s, \phi_s, q_s)$. The efficiency is the probability for a photon scattered into a direction (θ_s, ϕ_s) and polarization q_s to be detected. For directions and polarizations that the instrument does not collect, $\epsilon(\theta_s, \phi_s, q_s) = 0$. The scatter signal H measured by the system can then be expressed in terms of the BRDF as an efficiency-corrected directional-hemispherical reflectance[3]

$$H = \int dH = \sum\nolimits_{q_s} \int d\omega_s \cos\theta_s \ \mathrm{BRDF}(\theta_i, q_i; \theta_s, \phi_s, q_s) \ \epsilon(\theta_s, \phi_s, q_s), \tag{1}$$

where $d\omega_s = d\theta_s d\phi_s \sin\theta_s$ is the differential solid angle of the collected scattered light. The scatter signal is the radiant flux collected into the detection area, normalized by the radiant flux incident on the sample, and is therefore dimensionless. The scatter signal is sometimes referred to as haze by SSIS manufacturers.[4] Under the condition that the scattered light arises only from surface topography, which is sufficiently small to be handled by first-order perturbation theory, the BRDF is related to the power spectral density (PSD) of the surface roughness by[5–7]

$$\mathrm{BRDF}(\theta_i, q_i; \theta_s, \phi_s, q_s) = \frac{16\pi^2}{\lambda^4} \cos\theta_i \cos\theta_s \ Q(\theta_i, q_i, \theta_s, \phi_s, q_s) \ \mathrm{PSD}(f_x, f_y), \tag{2}$$

where $Q(\theta_i, q_i, \theta_s, \phi_s, q_s)$ is a polarization factor which depends upon the index of refraction of the material, and f_x and f_y are the components of the two-dimensional spatial frequency. We define the spatial frequency response function $\rho(f_x, f_y)$ of the scattered light collection system as that function which satisfies the relationship

$$H = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho(f_x, f_y) \mathrm{PSD}(f_x, f_y) df_x df_y. \tag{3}$$

Since many surfaces have isotropic power spectral densities, and some measurement systems spin the sample, it is appropriate to develop an isotropic spatial frequency response function. We begin by defining f and α so that

$$f_x = f\cos\alpha, \quad f_y = f\sin\alpha, \quad \text{and} \quad df_x df_y \to f \ df \ d\alpha. \tag{4}$$

Averaging over the rotational orientation of the sample is equivalent to averaging over α. Equation 3 can then be expressed as

$$\begin{aligned} H &= \int_0^{\infty} df \int_0^{2\pi} d\alpha \ f \ \rho(f, \alpha) \ \mathrm{PSD}(f) \\ &= \int_0^{\infty} df \ 2\pi f \ \mathrm{PSD}(f) \int_0^{2\pi} \frac{d\alpha}{2\pi} \rho(f, \alpha) \\ &= \int_0^{\infty} df \ 2\pi f \ \rho(f) \ \mathrm{PSD}(f), \end{aligned} \tag{5}$$

where

$$\rho(f) = \int_0^{2\pi} \frac{d\alpha}{2\pi} \rho(f, \alpha), \tag{6}$$

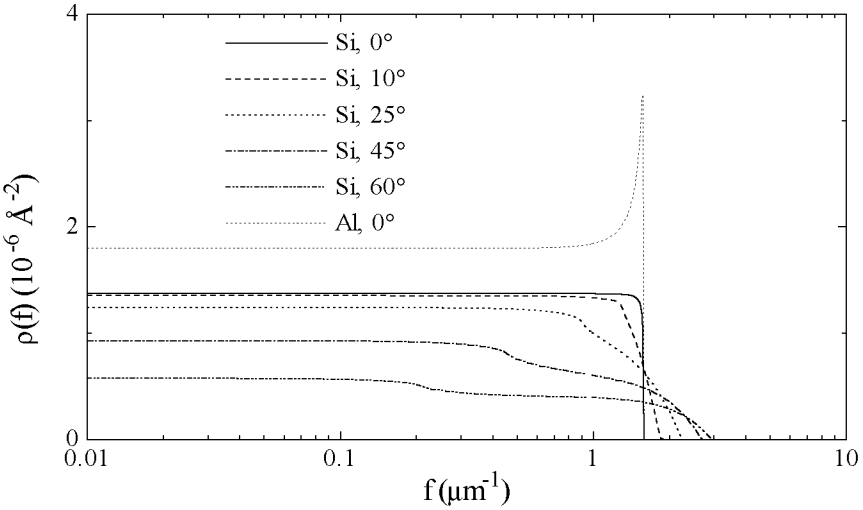

Figure 2 The spatial frequency response function calculated for the ideal hemispherical detection system with a variety of incident angles. Except where noted, the material is assumed to be silicon and the wavelength of the incident light is 632.8 nm.

and PSD(f) is the isotropic two-dimensional power spectral density evaluated at $f = (f_x^2 + f_y^2)^{1/2}$. We will now determine a procedure for calculating $\rho(f)$ for a variety of configurations. From the 2-d grating equations,

$$\begin{aligned} \lambda f_x &= \sin\theta_s \cos\phi_s - \sin\theta_i \\ \lambda f_y &= \sin\theta_s \sin\phi_s, \end{aligned} \tag{7}$$

it is straightforward to show that

$$d\omega_s = \sin\theta_s \; d\theta_s \; d\phi_s = \frac{\lambda^2}{\cos\theta_s} \; df_x \; df_y = \frac{\lambda^2}{\cos\theta_s} f \; df \; d\alpha. \tag{8}$$

Therefore,

$$\begin{aligned} dH &= \sum\nolimits_{q_s} d\omega_s \cos\theta_s \mathrm{BRDF}(\theta_i, q_i; \theta_s, \phi_s, q_s) \epsilon(\theta_s, \phi_s, q_s) \\ &= \sum\nolimits_{q_s} d\omega_s \frac{16\pi^2}{\lambda^4} \cos\theta_i \cos^2\theta_s \; Q(\theta_i, q_i, \theta_s, \phi_s, q_s) \; \mathrm{PSD}(f_x, f_y) \epsilon(\theta_s, \phi_s, q_s) \\ &= \sum\nolimits_{q_s} df \; d\alpha \; f \frac{16\pi^2}{\lambda^2} \cos\theta_i \cos\theta_s \; Q(\theta_i, q_i, \theta_s, \phi_s, q_s) \; \mathrm{PSD}(f_x, f_y) \epsilon(\theta_s, \phi_s, q_s) \\ &= \sum\nolimits_{q_s} df \; d\alpha \, f \frac{16\pi^2}{\lambda^2} \cos\theta_i \sqrt{1 - (\lambda f \cos\alpha + \sin\theta_i)^2 - (\lambda f \sin\alpha)^2} \, Q(\theta_i, q_i, \theta_s, \phi_s, q_s) \, \mathrm{PSD}(f_x, f_y) \epsilon(\theta_s, \phi_s, q_s), \end{aligned} \tag{9}$$

where we have also used the 2-d grating equation to eliminate θ_s. Comparing this result to Eqs. 5 and 6, it is apparent that

$$\rho(f) = \frac{8\pi}{\lambda^2} \cos\theta_i \int_0^{2\pi} d\alpha \sqrt{1 - (\lambda f \cos\alpha + \sin\theta_i)^2 - (\lambda f \sin\alpha)^2} \sum\nolimits_{q_s} Q(\theta_i, q_i, \theta_s, \phi_s, q_s) \; \epsilon(\theta_s, \phi_s, q_s). \tag{10}$$

The expression in Eq. 10 can be calculated numerically for any given geometry. The spatial frequency response function is written in a form that allows calculation at each spatial frequency f. During evaluation of the integral, conversion must be made from the variables f and α to the detection angles θ_s and ϕ_s in order to determine the detection efficiency $\epsilon(\theta_s, \phi_s, q_s)$ and polarization factor, $Q(\theta_i, q_i, \theta_s, \phi_s, q_s)$.

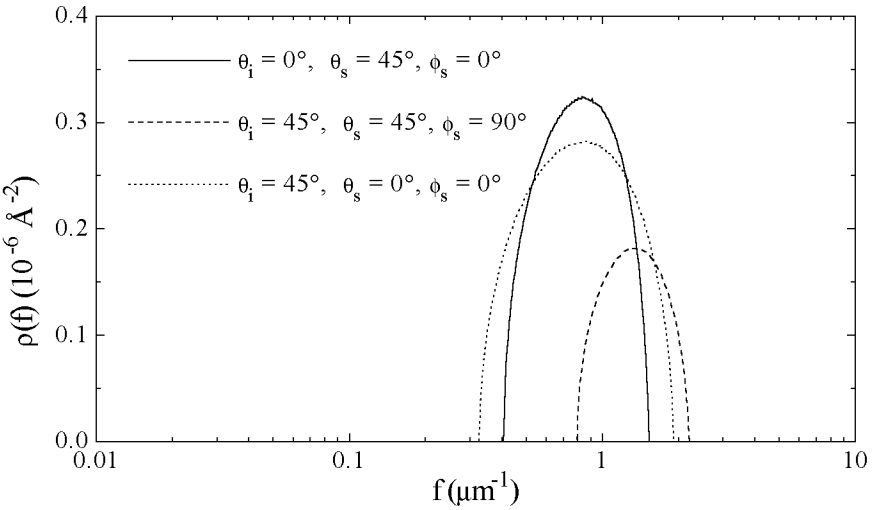

Figure 3 The spatial frequency response function calculated for a conical collection system for a variety of central incident/exitant angle combinations. The conical half-angle of collection is 30° for each curve. The material is assumed to be silicon and the wavelength of the incident light is 632.8 nm.

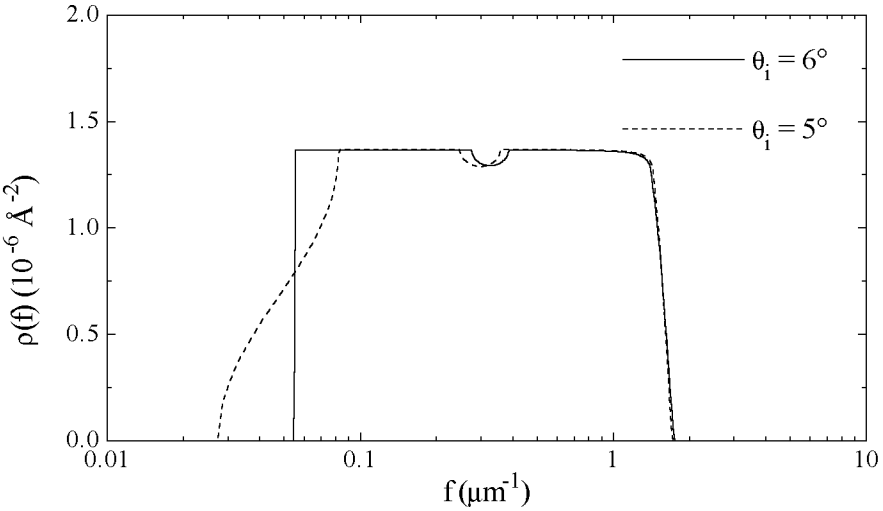

Figure 4 The spatial frequency response function calculated for a total integrated scatter collection system with input and output ports aligned at polar angles of 6° and having opening half-angles of 2°. The curves are shown for incident angles of 6° and 5°, illustrating the effects of misalignment. The material is assumed to be silicon and the wavelength of the incident light is 632.8 nm.

3. DISCUSSION

To illustrate, $\rho(f)$ has been calculated for a variety of simple detection systems. For simplicity, we have assumed the detection efficiency $\epsilon(\theta_s, \phi_s, q_s) = 1$ for all detected angles and $\epsilon(\theta_s, \phi_s, q_s) = 0$ otherwise. We have also assumed that the substrate is silicon ($n = 3.882 + 0.019i$), the input light is s-polarized, and the detection system is polarization insensitive ($Q = Q_{ss} + Q_{sp}$), and the wavelength is 632.8 nm. Figure 2 shows $\rho(f)$ for an ideal hemispherical detection system for a variety of input angles. This scheme is illustrative of the maximum attainable $\rho(f)$, but is physically unrealizable, since the incident light and the specular light have no ports to enter or exit the detection system. Figure 3 shows $\rho(f)$ for a conical detection scheme where the cone is centered about different angles. This scheme

is easily replicated in the laboratory by a collection lens focusing the point the laser strikes the sample onto the detector. Figure 4 shows the calculated $\rho(f)$ for a total integrated scatter (TIS) scheme, which collects all the light scattered into the hemisphere, except that which is scattered into entrance and exit ports. Figure 4 also illustrates the effect that a small misalignment on the input laser beam will have on the spatial frequency response function. Such a misalignment, in either direction, can be seen to shift the emphasis of the scatter signal toward lower spatial frequencies, and those for which the PSD is often dominant. Such changes suggest the need for having calibration mechanisms in place which allow the in-line measurement of $\rho(f)$.

The defined spatial frequency response function and the procedure for calculating it has a number of limitations that arise from the assumptions made during the derivation. In particular, the spatial frequency response function is only useful for characterizing the response of an instrument to an isotropic sample for which first-order perturbation theory is applicable. It does not address an optical scattering instrument's ability to detect non-topographic scatter such as particulate contamination, subsurface defects, or material inhomogeneities. Since these scattering mechanisms give rise to different angular dependences, their $Q(\theta_i, q_i, \theta_s, \phi_s, q_s)$ factors do not correspond to those for topographic scatter. It may be possible to define appropriate response functions for each of these other scatter mechanisms, but that topic is beyond the scope of this paper.

Due to the material dependence of $Q(\theta_i, q_i, \theta_s, \phi_s, q_s)$, the spatial frequency response function for a given optical geometry is dependent upon the material being examined. This is not a serious drawback, since many instruments are designed with a particular material in mind (such as silicon). To illustrate the material dependence, one of the curves shown for silicon is shown in Fig. 2 for aluminum[8] ($n = 1.39 + 7.65i$). Comparison of the results from calculations for aluminum and silicon reveals that the $\rho(f)$ differ, in a manner not correctable by a single constant. Furthermore, it has been found that a unique function does not exist which maps the aluminum $\rho(f)$ into the silicon $\rho(f)$. Therefore, if an instrument were to be designed for the measurement of various different materials, it would be important to specify the spatial frequency response function for each of those materials.

Another limitation of the spatial frequency response function definition is that it does not appropriately characterize systems which scan the wafer by scanning the incident laser beam, thus changing the incident angle at different spots on the wafer. Such an instrument would be characterized as having a non-uniform response function over the area of the sample. Such instrumentation can perhaps be characterized by specifying appropriate extreme examples of the response function at several spots on a wafer, or by specifying minimum and maximum values at each spatial frequency.

4. OPPORTUNITIES FOR REALIZATION

Section 2 of this paper presents a numerical approach to calculating the spatial frequency response function $\rho(f)$ of a particular instrument. In principle, such a calculation can serve as a starting point for specifying scattered light instrumentation. However, an ideal situation would arise if an artifact existed that allowed a particular instrument to be calibrated, yielding a particular spatial frequency response function for each instrument and allowing for periodic recalibration. Such an artifact may take the form of a wafer which carries a set of shallow sinusoidal gratings whose amplitude, spatial frequency, and grating directions are well characterized.[9,10] Appropriate measurement of the measured scattered light at each point on the test wafer would yield the information neccesary to extract an approximation to $\rho(f)$.

A similar, and probably more realistic, approach relies on having a set of N samples (possibly also patterned on a single wafer) which have PSDs that differ by something other than magnitude. For example, their PSDs may have different slopes. One approximation to the spatial frequency response function can be expressed as

$$\rho(f) = \sum_{m=1}^{M} r_m \rho_m(f), \tag{11}$$

where $\rho_m(f)$ is a set of M trial functions. Then, from Eq. 5, the measured scatter signal of the nth sample, H_n, will

be

$$H_n = \int_0^\infty 2\pi f P_n(f) \sum_{m=1}^{M} r_m \rho_m(f) df$$
$$= \sum_{m=1}^{M} r_m T_{nm}, \tag{12}$$

where

$$T_{nm} = \int_0^\infty 2\pi f P_n(f) \rho_m(f) df, \tag{13}$$

and $P_n(f)$ is the PSD of the nth sample. Then we can determine r_m, and therefore $\rho(f)$, from

$$r_m = \sum_{n=1}^{N} T_{nm}^{-1} H_n. \tag{14}$$

Under the conditions of the grating standard mentioned above, the trial functions $\rho_m(f)$ would be top-hat functions and the PSDs would be δ functions. Therefore, the matrix T_{nm} is diagonal and, thus, invertible. In the more general case, the P_n and ρ_m have to be linearly independent in order for T_{nm} to be invertible (although this is not a sufficient condition). Samples with PSDs having different slopes (when plotted on log-log plots) would satisfy this requirement, although other arrangements are possible.

5. SUMMARY

This paper presents a methodology by which low-level optical scatter measurement systems can be characterized and calibrated. This methodology is based upon a definition of a spatial frequency response function $\rho(f)$, which characterizes the sensitivity of the instrument to a particular spatial frequency of surface roughness.

6. REFERENCES

[1] W. M. Bullis, *Microroughness of Silicon Wafers*, Semiconductor Silicon/1994, H. R. Huff, W. Bergholz, and K. Sumino, ed. ESC Proceedings, **94-10**, pp. 1156-69

[2] E. Church, G. Sanger, and P. Takacs, *Comparison of Wyko and TIS measurements of surface finish*, SPIE Proceedings **749**, 65 (1987).

[3] F. E. Nicodemus, J. C. Richmond, J. J. Hsia, I. W. Ginsberg, and T. Limperis, *Geometrical Considerations and Nomenclature for Reflectance*, NBS Monograph 160 (National Bureau of Standards, Gaithersburg, 1977)

[4] We use the term *scatter signal* since its definition differs from that of *haze* as outlined by the ASTM Standard E284-94a, Standard Terminology of Appearance, (American Society for Testing and Materials, Philadelphia, 1991).

[5] D. E. Barrick, *Radar Cross Section Handbook*, (Plenum,New York,1970).

[6] S. O. Rice, *Reflection of Electromagnetic Waves from Slightly Rough Surfaces*, Comm. Pure and Appl. Math. **4**, 351-78 (1951).

[7] J. C. Stover, *Optical Scattering: Measurement and Analysis*, (McGraw-Hill,New York,1990).

[8] Aluminum has been shown in the past to not follow scaling laws appropriate for topographic scattering; see e.g. Ref. 7. The use of aluminum as an example is purely illustrative.

[9] E. Marx, T. R. Lettieri, T. V. Vorburger, and M. McIntosh, *Sinusoidal surfaces as standards for BRDF instruments*, Proc. SPIE **1530**, 15-21 (1991).

[10] E. Marx, T. R. Lettieri, and T. V. Vorburger, *Light scattering by sinusoidal surfaces: illumination windows and harmonics in standards*, Appl. Opt. **34**, 1269-77 (1995).